

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 November 2000 (30.11.2000)

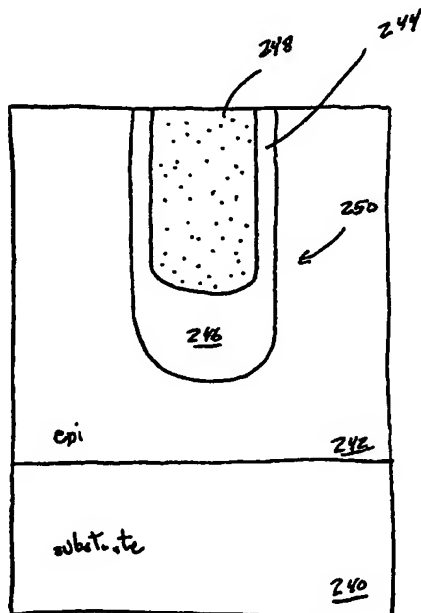
PCT

(10) International Publication Number
WO 00/72372 A1

- (51) International Patent Classification⁷: H01L 21/336 (74) Agents: STEUBER, David, E. et al.; Skjerven, Morrill, MacPherson, Franklin & Friel LLP, Suite 700, 25 Metro Drive, San Jose, CA 95110 (US).
- (21) International Application Number: PCT/US00/14363
- (22) International Filing Date: 24 May 2000 (24.05.2000) (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/318,403 25 May 1999 (25.05.1999) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant: WILLIAMS, Richard, K. [US/US]; 10292 Norwich Avenue, Cupertino, CA 95014 (US).
- (71) Applicant and
(72) Inventor: GRABOWSKI, Wayne, B. [US/US]; 1390 Miravalle Avenue, Los Altos, CA 94024 (US). Published:
— With international search report.

[Continued on next page]

(54) Title: TRENCH SEMICONDUCTOR DEVICE HAVING GATE OXIDE LAYER WITH MULTIPLE THICKNESSES AND PROCESSES OF FABRICATING THE SAME.



(57) Abstract: A trench semiconductor device such as a power MOSFET the high electric field at the corner of the trench (250) is diminished by increasing the thickness of the gate oxide layer (244) at the bottom of the trench (250). Several processes for manufacturing such devices are described. In one group of processes a directional deposition of silicon oxide (272) is performed after a trench (268) has been etched, yielding a thick oxide layer (270) at the bottom of the trench (268). Any oxide which deposits on the walls of the trench (268) is removed before a thin gate oxide layer (276) is grown on the walls. The trench (268) is then filled with polysilicon (278) in one or more stages. In a variation of the process a small amount of photoresist (310) is deposited on an oxide (270) at the bottom of the trench (268) before the walls of the trench (268) are etched. Alternatively, polysilicon (320) can be deposited in a trench (268) and etched back until only a portion (322) remains at the bottom of the trench (268). The polysilicon (320) is then oxidized and the trench (268) is filled with polysilicon. The processes can be combined, with a directional deposition of oxide being followed by a filling and oxidation of polysilicon. A process of forming a "keyhole" shaped gate electrode (634) includes depositing polysilicon at the bottom of the trench (606), oxidizing the top surface of the polysilicon, etching the oxidized polysilicon, and filling the trench (606) with polysilicon.

WO 00/72372 A1



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**TRENCH SEMICONDUCTOR DEVICE HAVING GATE OXIDE LAYER
WITH MULTIPLE THICKNESSES AND PROCESSES
OF FABRICATING THE SAME**

5

FIELD OF THE INVENTION

This invention relates to semiconductor devices having a gate electrode that
10 is embedded in a trench and in particular to structures and methods of protecting
such devices against damage to the gate oxide layer when the devices is subjected
to high voltage differences while in an off condition. The invention particularly
relates to trench MOSFETs.

BACKGROUND OF THE INVENTION

15 There is a class of semiconductor devices in which a gate electrode is
formed in a trench that extends from the surface of a semiconductor chip. One
example is a trench-gated MOSFET, and other examples include insulated gate
bipolar transistors (IGBTs), junction field-effect transistors (JFETs) and
accumulation-mode field-effect transistors (ACCUFETs). All of these devices
20 share the common characteristic of a trench structure where the bottom of the
trench for some reason can be exposed to high electric fields or where the bottom
of the trench might form a parasitic capacitor including the gate electrode and the
semiconductor material surrounding the trench.

Figs. 1 through 10 show cross-sectional views and characteristics of known
25 trench-gated devices. Fig. 1 shows a trench-gated MOSFET 100 having a top
metal layer 102, a gate 104 formed in a trench 106 and separated from an epitaxial
silicon layer 108 by a gate oxide layer 110. MOSFET 100 also includes an N+
source region 112 and a P-body 114. The drain of MOSFET 100 includes the N-

epi layer 108 and an N+ substrate 116. A deep P+ region 118 is created under P-body 114, as suggested in U.S. Patent No. 5,072,266 to Bulucea et al. The PN junction between deep P+ region 118 and N-epi layer 108 forms a voltage-clamping diode 117 where avalanche breakdown normally occurs. A P+ body
5 contact region 119 forms a contact between metal layer 102 and P-body 114. The gate, which is typically formed of polysilicon, is protected from the metal layer 102 by an oxide layer 120 that is above the gate 104 and that is patterned with a feature that does not correspond to the trench itself, typically a contact mask.

As shown, gate oxide layer 110 consists of a uniform thin layer of oxide
10 along the three sides of the polysilicon gate 104. That is, the portions of gate oxide layer 110 on the sidewalls of the trench and also the curved and linear portions of the gate oxide layer 110 at the bottom of the trench (except for some stress-related and etch-related changes in the oxide thickness that occur at the trench bottom) are generally of a uniform thickness in the range of, for example, 150 Å to 1,200 Å.

15 There are many variations of this general type of MOSFET. For example, Fig. 2 shows a MOSFET 130 which is generally similar to MOSFET 100 but does not include a deep P+ region 118. The gate of MOSFET 130 protrudes slightly through P-body 132 because the depth of P-body 132 and the depth of the trench 134 are determined in two unrelated processes. Thus, in vertical devices there is
20 no guarantee of the net overlap of the polysilicon gate into the drain region. It turns out that this variation affects the operation of the device and may affect its reliability as well. Also, in Fig. 2 there is no additional diode formed by the deep P+ region 118 to clamp the voltage, so breakdown can occur wherever the voltage is raised to the point that the device goes into avalanche.

25 MOSFET 140, shown in Fig. 3, is variation of MOSFETs 100 and 130, where the MOSFET cells 142 contain no deep P+ region, and a diode cell 144 containing a deep P+ region is distributed at predetermined intervals throughout the array to act as a voltage clamp and limit the strength of the electric fields in the MOSFET cells. In MOSFET 140, the gate oxide layer is of uniform thickness.

Figs. 4A-4G illustrate various aspects of the breakdown phenomenon. Fig. 4A shows the electric field strength contours at breakdown in a trench-gated device 150 having a relatively thick gate oxide layer. Device 150 is in effect a gated diode, a structural element of most trench-gated vertical power MOSFETs. As indicated, the strongest electric field, where impact ionization would occur during avalanche breakdown, is located at the junction directly beneath the P+ body region. In contrast, device 160, shown in Fig. 4B, has a relatively thin gate oxide layer. While some ionization still occurs underneath the P+ region, the highest electric field levels are now located near the corner of the trench. A field plate induced breakdown mechanism causes the strength of the electric field to increase.

Figs. 4C and 4D show the ionization contours of devices 150 and 160, respectively, when they go into avalanche breakdown. Whether there is a thick gate oxide layer, as in Fig. 4C, or thin gate oxide layer, as in Fig. 4D, eventually in "deep" avalanche, i.e., when the device is forced to conduct large currents in avalanche, breakdown starts to occur at the corner of the trench. Even in the thick oxide case (Fig. 4C), where the peak electric field is not at the corner of the trench (Fig. 4A), as the drain voltage increases eventually ionization occurs at the corner of the trench. However, there are more contours in Fig. 4D, indicating a higher ionization rate where the gate oxide layer is thin.

Fig. 4E shows that if one introduces a diode clamp including a deep P+ region, as shown on the right-hand side, the diode will break down at a lower voltage, and avalanche breakdown should not occur at the corner of the trench. If the resistance of the current path through the diode is low enough, then the diode will clamp the maximum voltage of the device. As a result, the voltage will never rise to the point that avalanche breakdown occurs near the corners of the trenches.

Fig. 4F is a graph showing the breakdown voltage (BV) as a function of gate oxide thickness (X_{OX}) for 20 V and 30 V devices. The doping concentration of the epitaxial (epi) layer in the 30 volt device is more lightly doped. The 30 V device would ideally have an avalanche breakdown of around 38 volts. In the 20 volt device the epi would be more heavily doped and the device would ideally have

an avalanche breakdown of around 26 or 27 V. As the gate oxide is thinned from 1,000Å to a few hundred Å, basically the breakdown voltages are relatively constant or may actually even increase somewhat as the shape of the field plate of the gate is actually beginning to help relax the electric field. At thicknesses of less
5 a few hundred Å, however, breakdown degradation begins to occur.

Beyond the point where the breakdown voltage begins to drop (below 30 V for the 30 V device epi and below 20 V for the 20 V device) is the area labeled field plate induced (fpi) breakdown. In this area, breakdown occurs near the trench. For a reliable device one needs to add a diode clamp having a breakdown
10 that is lower than the breakdown in the field plate induced area, so that the diode breaks down first. With a diode having a breakdown voltage as shown in Fig. 4F, breakdown would never occur near the gate in the 30 V device, but that diode would have too high a breakdown voltage to protect a 20 V device. To protect the 20 V device, the breakdown voltage of the diode clamp would have to be below the
15 curve for the 20 V device.

Fig. 4G is a schematic diagram of the devices shown in Figs. 4A-4D showing a gated diode in parallel with a MOSFET and a diode voltage clamp in parallel with both the MOSFET and gated diode. The arrangement is designed such that the diode clamp breaks down first. The gated diode never "avalanches"
20 before the diode clamp. This becomes more and more difficult to do as the gate oxide layer becomes thinner.

Figs. 5A and 5B show the ionization contours in a device 170 having a sharp trench corner and a device 172 having a rounded trench corner. Fig. 5B indicates that rounding the trench corners does reduce the magnitude of the
25 ionization, but ultimately if one drives the device deeply enough into breakdown, the breakdown still occurs at the trench corner, and the device is at risk.

Figs. 6A-6C show the electric field strength contours, the equipotential lines and the electric field lines, respectively, in a MOSFET 180. The gate of MOSFET 180 is tied to the source and body and is grounded, and the drain is
30 biased at V_D . From Fig. 6B it is evident that the drain voltage V_D is divided and

spaced out across the region. On the left hand side of Fig. 6B, the equipotential lines are squeezed closer together, and particularly around the trench corner they are squeezed even tighter. This produces electric field lines that are at right angles to the equipotential lines, as shown in Fig. 6C. One can see why a high electric
5 field occurs at the trench corner and why rounding the corner does not solve this problem. It is basically a volumetric problem in that there is an electric field that terminates on an electrode having a lower surface area, namely the gate, and so the electric field lines are crowded at the corner.

Fig. 6D shows MOSFET 180 when it is turned on by putting a positive
10 voltage V_G on the gate. A current flows down the side wall of trench and then it also spreads out along the bottom of the trench and into the region below the mesa at an angle from the side of the trench. However, in the process the current flows through areas that have high electric fields, as shown by the electric field contours of Fig. 6A. When a high current flows through an area that has a high field (and
15 that would be the case where the device is saturated), the current carriers collide with the atoms of the epi layer and knock off, by momentum transfer, additional carriers. This forms new electron-hole pairs that in turn are accelerated and create additional collisions, ionizing additional atoms.

Fig. 6E shows the ionization contours in MOSFET 180 when it is in the on
20 state. The ionization contours shown in Fig. 6E are different from those shown in Fig. 4C, for example, when device 150 is in the off state. The difference is that the ionization contours pull upwards all the way around the side of the trench, even up near the P-body. This has a number of damaging effects on the device. One effect is that it creates electron-hole pairs in the vicinity of the gate oxide that can be
25 accelerated quite easily by the high electric field in that area. The electron-hole pairs can actually be trapped in the gate oxide, and they can damage the gate oxide.

Furthermore, this phenomenon produces an upper limit in the amount of voltage that one can put on the device, because so many electron-hole pairs may be produced that they begin to modulate the effective doping concentration of the
30 epitaxial layer, by making the region around the side of the trench seem more

heavily doped than it actually is. That occurs because electrons from the newly generated electron-hole pairs are swept into the substrate by the positive drain voltage V_D , and the holes are swept into the P-body. The net effect is that, since the electrons and holes can only travel at a certain velocity, the local charge distribution adjusts itself to maintain charge neutrality. Specifically, surrounding the reverse-biased junction is a region known as a depletion region or space charge region, where (in the absence of impact ionization) no free charge carriers are present. The immobile charge residing within the depletion region, namely positive ions on the N-type side of the junction and negative ions on the P-type side of the junction, produces a "built-in" electric field across the junction. In the presence of impact ionization, the holes drifting across the N-type region add to the positive fixed charge and thereby increase the electric field, further enhancing the impact ionization process. These excess holes make the epitaxial region, which in this example is N-type material, appear more heavily doped because of the increase in the "built-in" field. The net effect is an increase in the electric field, which degrades the breakdown. This effect is shown in the current-voltage characteristics of Fig. 6F where the drain current I_D increases dramatically at a certain drain voltage. The drain voltage at which this happens is the same for each of the gate voltages shown. This problem becomes worse as the gate oxide is thinned.

Another problem with the trench device relates to capacitance. Fig. 7A is a schematic diagram of a MOSFET 190 having a gate driven by a current source 192 and having resistive load 194. A voltage source 196 connected to the source and drain supplies a voltage V_{DD} resulting in a drain voltage V_D at the drain. As shown in Figs. 7B-7D, at a time t_1 current source 192 begins to supply a constant current to the gate and the voltage on the gate relative to the source, labeled V_G in Fig. 7C, starts to rise. But because it does not immediately hit threshold, the drain voltage V_D does not start to fall because MOSFET 190 is not yet turned on. As soon as the V_G hits threshold, at time t_2 , MOSFET 192 saturates and turns on and carries current. V_D starts to drop, but as it starts to drop it causes a capacitive coupling between the drain and the gate of MOSFET 192 and halts the upwards progression

of the gate voltage V_G . V_G remains flat until MOSFET 192 gets into its linear region. Then, MOSFET 192 begins to look like an on-resistance in a voltage divider, with a small voltage across MOSFET 192 and most of the voltage V_{DD} across resistor 194.

5 At that point the capacitive coupling effect between gate and drain is satisfied and the V_G continues its progress to a higher voltage. The plateau is due to a gate-to-drain overlap capacitance similar to the Miller effect, but this is not a small signal effect. This is a large signal effect. At that time the drain current I_D also continues to rise, but as shown in Fig. 7D its upward progression is slowed.

10 Fig. 7E shows a plot of V_G as function of the charge on the gate Q_G , where Q_G is equal to I_G times the time t , I_G being a constant. The gate voltage V_G rises to a certain level, then it remains constant, and then it rises again. If there were no feedback capacitance between the drain and gate, the voltage would rise linearly, but instead the straight line is interrupted by the plateau.

15 In Fig. 7E, the point V_{G1} , Q_{G1} corresponds to a certain capacitance because C is equal to ΔQ over ΔV . Since it takes more charge to get to the point, Q_{G2} and V_{G1} , then that point reflects more capacitance. So the capacitance in the device, as shown in Fig. 7F, starts at a low value C_{ISS} , which is relatively constant, and then it jumps to a higher effective value $C_G(\text{eff})$, and then it is relatively constant.

20 Because of this effect the device has a higher effective capacitance than is desirable during the switching transition. As a result, there is an undue amount of energy lost in turning the device on.

 As shown in Fig. 7G, the input capacitance actually has a number of components, including the gate-to-source capacitance C_{GS} and the gate-to-body capacitance C_{GB} , neither of which exhibits the amplification effect of the gate-to-drain capacitance C_{GD} . The gate-to-drain capacitance C_{GD} is shown in Fig. 7G, around the bottom and side wall of the trench. The equivalent schematic is shown in Fig. 7H. Even if C_{GD} is the same order of magnitude as C_{GS} and C_{GB} ,

25

electrically it will look much larger (e.g., 5 to 10 times larger) because it is amplified during the switching process.

As indicated above, rounding the trench bottom helps to limit the damage to the gate oxide layer, although it is not a complete solution to the problem. Figs. 8A-8C illustrate a process for forming a trench with rounded corners. In Fig. 8A small reaction ions 202 etch the silicon through an opening in a mask 200 at the surface. Ions 200 are accelerated by an electric field in a downward direction such that they etch a trench having essentially a straight side wall. When the trench reaches a certain depth the electric field is relaxed, as shown in Fig. 8B. Alternatively, one could possibly change the chemistry. At the end of the process, as shown in Fig. 8C, the electric field is modified so that the etching ions are traveling in all different directions. That begins to not only widen the trench, but also rounds out the bottom. Hence, the process includes an anisotropic etch that is converted to an isotropic etch. The anisotropy is also influenced by the formation of a polymer as a by-product of the etching operation on the sidewall of the trench. If the chemistry removes the polymer as soon as it forms, the etch will behave in a more isotropic way. If the polymer remains on the sidewall, only the bottom of the trench will continue to etch.

Figs. 9A-9D show a method that includes creating a mask 210 (Fig. 9A), etching the trench 212 (Fig. 9B), forming an oxide layer 214 on the walls of the trench (Fig. 9C), which may be removed and then re-grown to remove defects (this is called sacrificial oxidation), and then filling the trench with a polysilicon layer 216 (Fig. 9D).

Figs. 10A-10F illustrate a typical process of forming a trench MOSFET. The process starts with an N-epitaxial layer 220 grown on an N⁺ substrate 222 (Fig. 10A). Using the process of Figs. 9A-9C, for example, a polysilicon-filled trench 224 is formed in N-epi layer 220 (Fig. 10B). The surface may or may not be planar depending on how the surface oxides are made in the process. Then a P-body 226 is introduced, although the P-body 226 could be introduced prior to the formation of the trench 224 (Fig. 10C). Both process flows are manufacturable,

but forming the trench first is preferable because the etching process can influence the doping concentrations in the P-body. Then the surface is masked and an N+ source region 228 is implanted (Fig. 10D). An optional shallow P+ region 230 is implanted to ohmic contact between the P-body and a metal layer to be deposited later. P+ region 230 can be implanted through an opening in an oxide layer 232 that is deposited across the region and then etched to form a contact mask (Fig. 10E). The contact mask may or may not be used to define the P+ region 232. Finally, a metal layer 234 is deposited on the surface to contact the N+ source region 228 and P+ region 230 (Fig. 10F).

10 SUMMARY OF THE INVENTION

In accordance with this invention, a trench-gated semiconductor device is formed, having a dielectric layer separating the gate electrode from the semiconductor material surrounding the trench wherein the thickness of the dielectric layer is greater in a region at the bottom of the trench. This structure helps to reduce the strength of the electric field near the bottom of the trench, particularly at the corner or rounded portion where the bottom of the trench makes a transition to a sidewall of the trench, and to reduce capacitance.

Several processes are used to fabricate this structure. One process includes the following steps. A trench is etched in the semiconductor material. A directional deposition of a dielectric material is then performed such that the dielectric material is deposited preferentially on horizontal surfaces such as the bottom of the trench. This is done by creating an electric field in the deposition chamber (e.g., a chemical vapor deposition or sputtering chamber) so as to accelerate the charged ions of the dielectric towards the semiconductor material. The trench is filled with a conductive material that will form the gate electrode. Following the directional deposition, any of the dielectric that was deposited on the sidewall of the trench can be removed, and a convention dielectric layer can be grown on the sidewall of the trench. In many processes, the dielectric material is silicon dioxide and the conductive material is polysilicon.

In one process, the conductive material is etched back to a level roughly coplanar with the surface of the semiconductor material, and a dielectric layer is deposited over the top surface of the dielectric material. In one variant, the conductive material (e.g., polysilicon) is oxidized to form an oxide layer, preferably after the conductive material has been etched back into the trench. The conductive material can be oxidized to a thickness such that the oxide itself is adequate to insulate the gate electrode, or another conductive material, such as glass, can be deposited over the oxidized conductive material.

In another variant, the conductive material that forms the gate electrode is deposited in two stages.

In another alternative, a masking material such as photoresist is applied after the preferential deposition of the dielectric material. The masking material is removed from all locations except the bottom of the trench, and the trench is subjected to an etch or dip to remove dielectric material from the sidewalls of the trench. A dielectric layer is then formed on the sidewalls of the trench.

In yet another alternative, following the directional deposition of the dielectric, a material such as polysilicon that can be oxidized to form a dielectric is deposited and etched back until only a portion of the material remains on top of the dielectric at the bottom of the trench. The material is then oxidized to form a thicker dielectric layer at the bottom of the trench.

Another group of alternatives avoids the directional deposition of a dielectric material. Instead a material such as polysilicon that can be oxidized to form a dielectric is deposited and etched back until only a portion remains at the bottom of the trench.

Processes in accordance with this invention may include a process for self-aligning the trench with a contact to the top surface of the "mesa" between the trenches. A "hard" layer of a material such as silicon nitride is used as a trench mask. The hard mask remains in place until a dielectric layer has been formed over the gate electrode, preferably by oxidizing polysilicon gate. The hard mask is then

removed, exposing the entire top surface of the mesa and allowing a contact to be made thereto with a metal layer.

5 A process of this invention may include the use of a sidewall spacer near the top corners of the trench to prevent a short between the gate electrode and the semiconductor mesa. After the trench mask has been deposited and an opening defining the location of the trench has been made in the trench mask, a layer of a "hard" material such as silicon nitride, and optionally an overlying oxide, is isotropically deposited into the opening in the trench mask. The "hard" material is deposited on the exposed edges of the trench mask. An etch is then performed,
10 following which the surface of the semiconductor material is exposed in the central region of the opening but some of the deposited dielectric remains on the side edges of the trench mask, forming sidewall spacers. The trench is then etched. The dielectric sidewall spacers provide additional insulation between the later formed gate electrode and the semiconductor material in the mesa.

15 Another group of processes provide a "keyhole" shaped trench, wherein a thick dielectric layer extends some distance upward on the sidewalls of the trench. After the trench has been etched, a relatively thick oxide lining is grown or deposited on the bottom and sidewalls of the trench. The trench is filled with polysilicon, and the polysilicon is then etched back such that only a portion
20 remains at the bottom of the trench, overlying the oxide lining. The exposed oxide lining is removed from the sidewalls of the trench. The polysilicon is then partially oxidized by heating to form an oxide layer at its exposed surface, and during the same heating process an oxide layer is formed on the sidewalls of the trench. The trench is then subjected to an oxide etch, which removes the oxide layer formed
25 from the polysilicon as well as some of the oxide layer from the sidewalls of the trench. The trench is then refilled with polysilicon to yield a keyhole-shaped gate electrode.

In a variant of the above process for forming a keyhole-shaped gate electrode, after the oxide lining has been formed on the bottom and sidewalls of the
30 trench, an amount of a masking material such as photoresist is deposited over the

oxide lining at the bottom of the trench. An oxide etch is then performed to removed the oxide lining from the sidewalls of the trench, and the masking material is removed from the bottom of the trench. A relatively thin gate oxide layer is grown on the sidewalls of the trench, and the trench is filled with a
5 conductive material such as polysilicon which forms the gate electrode.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a cross-sectional view of a prior art trench power MOSFET having a deep P+ diode which functions as a voltage clamp.

Fig. 2 is a cross-sectional view of a prior art trench power MOSFET having
10 a flat body-drain junction.

Fig. 3 is a cross-sectional view of a prior art trench power MOSFET having a voltage clamp which is distributed among MOSFET cells which contain a flat body-drain junction.

Fig. 4A is a cross-sectional view showing the electric field contours in a
15 MOSFET having a thick gate oxide layer.

Fig. 4B is a cross-sectional view showing the electric field contours in a MOSFET having a thin gate oxide layer.

Fig. 4C is a cross-sectional view showing the ionization contours in a MOSFET having a thick gate oxide layer at the onset of avalanche breakdown.

Fig. 4D is cross-sectional view showing the ionization contours in a
20 MOSFET having a thin gate oxide layer at the onset of avalanche breakdown.

Fig. 4E is a cross-sectional view showing the ionization contours in a device which contains a deep P+ region used as a voltage clamp.

Fig. 4F is a graph showing the breakdown voltage as a function of gate
25 oxide thickness in MOSFETs fabricated in epitaxial layers having different doping concentrations.

Fig. 4G is a schematic diagram of a trench power MOSFET with an anti-parallel diode clamp.

Fig. 5A is a cross-sectional view showing ionization contours in a trench power MOSFET having a square trench corner.

Fig. 5B is a cross-sectional view showing ionization contours in a trench power MOSFET having a rounded trench corner.

5 Fig. 6A is a cross-sectional view showing the electric field contours in a trench power MOSFET having a flat body-drain junction.

Fig. 6B is a cross-sectional view showing the equipotential lines in a trench power MOSFET having a flat body-drain junction.

10 Fig. 6C is a cross-sectional view showing the electric field lines in a trench power MOSFET having a flat body-drain junction.

Fig. 6D is a cross-sectional view showing the current flow lines in a trench power MOSFET having a flat body-drain junction.

Fig. 6E is a cross-sectional view showing the ionization contours in a trench power MOSFET when it is turned on.

15 Fig. 6F is a graph showing a family of I-V curves for a power MOSFET at different gate voltages, showing how the sustaining voltage is reduced by impact ionization.

Fig. 7A is a schematic diagram of a gate charging circuit for a power MOSFET.

20 Fig. 7B is a graph illustrating the step function application of a gate drive current to a power MOSFET.

Fig. 7C is a graph illustrating how the gate voltage and drain voltage vary under the conditions of Fig. 7B.

25 Fig. 7D is a graph showing how the drain current varies under the conditions of Fig. 7B.

Fig. 7E is a graph showing how the gate voltage varies as a function of charge.

Fig. 7F is a graph showing how the effective input capacitance varies as a power MOSFET is turned on.

Fig. 7G is a cross-sectional view showing the components of the gate capacitance in a trench power MOSFET.

5 Fig. 7H is an equivalent circuit diagram of a trench MOSFET showing the inter-electrode capacitance.

Fig. 8A-8C are cross-sectional views showing how a gate trench having rounded corners is formed.

10 Figs. 9A-9D are cross-sectional views showing a process of etching a gate trench and filling the trench with polysilicon.

Figs. 10A-10F are cross-sectional views of a process of fabricating a conventional trench power MOSFET.

Fig. 11A is a cross-sectional view of a trench power MOSFET having a thick oxide layer at the bottom of the trench.

15 Fig. 11B is a cross-sectional view showing the MOSFET of Fig. 11A having a thick oxide layer patterned on the top surface of the semiconductor.

Fig. 11C is a cross-sectional view of the power MOSFET of Fig. 11A with a thick overlying oxide layer that is aligned to the walls of the trench.

20 Fig. 12 is a schematic flow diagram showing a number of process sequences in accordance with this invention.

Figs. 13A-13N illustrate a process sequence for fabricating a trench power MOSFET having a thick oxide layer at the bottom of the trench, using a directional deposition of an oxide layer and etching the polysilicon to a level even with the top of the semiconductor material.

25 Figs. 14A-14F illustrate an alternative process sequence in which the polysilicon is etched to a level below the surface of the semiconductor material and then oxidized.

Figs. 15A-15F illustrate an alternative process sequence in which the polysilicon is deposited in two stages.

Figs. 16A-16E illustrate an alternative process in which a small amount of photoresist is used to mask the thick oxide at the bottom of the trench.

5 Figs. 17A-17F illustrate a process in which the polysilicon is etched to a level near the bottom of the trench and then oxidized.

Figs. 18A-18F illustrate an alternative process in which the polysilicon is oxidized.

10 Figs. 19A-19L illustrate a process of fabricating a trench power MOSFET having an oxide layer over the gate electrode which is self-aligned with the walls of the trench.

Figs. 20A-20F illustrate a process sequence for fabricating a trench gate in an active array portion of a power MOSFET as well as a gate bus.

15 Figs. 21A-21E illustrate a problem that can occur from undercutting the thin oxide layer below the nitride.

Figs. 22A-22C illustrate further examples of this problem.

Figs. 23A-23G illustrate other problems that can arise in the fabrication of power MOSFETs in accordance with this invention.

20 Figs. 24A-24F illustrate problems that can occur from undercutting a hard mask during the removal of the top oxide in a self-aligned device.

Figs. 25A-25H illustrate a process of fabricating a trench power MOSFET with a thick bottom oxide layer and a nitride side spacer.

Figs. 26A and 26B illustrate a problem that can occur during the formation of the gate oxide layer in a thick bottom oxide device.

25 Figs. 27A-27D illustrate a method of avoiding the problem illustrated in Figs. 26A and 26B.

Figs. 28-33 illustrate different types of trench power MOSFETs that can be fabricated in accordance with this invention.

Fig. 34 illustrates a flow diagram of a process sequence of fabricating a trench power MOSFET using a conventional contact mask and incorporating a thick bottom oxide layer.

Figs. 35A-35L illustrate cross-sectional views showing the process of Fig. 34.

Figs. 36-39 are cross-sectional views showing trench power MOSFETs having "keyhole" shaped gate electrodes.

Figs. 40A-40L illustrate a process sequence for fabricating a MOSFET having a keyhole-shaped gate electrode.

Figs. 41A-41F illustrate an alternative process sequence of fabricating a MOSFET having a keyhole-shaped gate electrode.

Figs. 42A-42C illustrate the strength of the electric field in a conventional power MOSFET, a power MOSFET having a thick bottom gate oxide, and a power MOSFET having a keyhole-shaped gate electrode, respectively.

DESCRIPTION OF THE INVENTION

The problems associated with interactions between the gate and the drain of a MOSFET can be solved in part by reducing the coupling capacitance between them. In accordance with this invention, this is done by thickening the gate oxide layer at the bottom of the trench. Figs. 11-27 show various structures and sequences for forming a thick gate oxide on the bottom of the trench.

Fig. 11A shows an epitaxial ("epi") layer 242 grown on a substrate 240. A trench 250 is formed in epi layer 242. A gate oxide layer 244 lines the walls of trench 250, and a thick portion 246 of gate oxide layer 244 is located at the bottom of trench 250. Trench 250 is filled with polysilicon 248. Note that there is no oxide layer on top of polysilicon 248. The arrangement of Fig. 11A could be an intermediate structure; an oxide layer could be formed on top of polysilicon 248 at

a later stage of the process. Polysilicon 248 is typically doped to a heavy doping concentration. It may be formed with a top surface substantially planar, i.e., flat, with the silicon epi surface by a number of means. One method to make the surface flat is to deposit the polysilicon layer to a greater thickness and then etch it back.

- 5 Another means to produce a flat surface is to deposit the polysilicon to a thickness greater than the amount needed to fill the trench and then chemical mechanically polish the surface flat. A flat surface is desirable to reduce the height of steps which may form later in the fabrication process.

Fig. 11B shows a structure with an oxide layer 252 on top of polysilicon
10 layer 248. Since the lateral edges of oxide layer 252 do not correspond to the walls of trench 250, oxide layer 252 is most likely formed with a mask and etching step. Oxide layer 252 could be either deposited (e.g., by chemical vapor deposition) or it could be thermally grown or some combination of these steps. Fig. 11C shows a top oxide layer 254 that is grown in accordance with the teachings of Application
15 No. 09/296,959, which is incorporated herein by reference in its entirety. The sides of oxide layer 254 are generally aligned with the walls of trench 250 and oxide layer 254 extends down into trench 250. Polysilicon layer 248 is thus embedded in trench 250. The embodiments of Figs. 11B and 11C both have a thick gate oxide region 246 at the bottom of the trench.

- 20 Fig. 12 is a schematic diagram of several process flows that can be used to fabricate gate trenches in accordance with this invention. The details of these process flows are shown in Figs. 13-20. Fig. 12 illustrates in block diagram form that the trench may be formed using a photoresist mask or a hard mask sequence, followed by a directed oxide deposition planarized by either a selective etch, a
25 dipback, or a selective oxidization. A selective oxidization can be used without a directed deposition. Finally, the trench is filled with polysilicon using a one-step or two-step process.

More specifically, starting at the left side of Fig. 12, there are two options for forming the trench. In one option, shown in Figs. 13-18, the trench is formed
30 using a mask that is later removed, so that the mask is not available as a reference

for other processing steps. The other option is to use a "hard" mask to form the trench, as described in the above-referenced Application No. 09/296,959, which is then employed as a reference later in the process. This option is generally described in Figs. 19 and 20. After the trench is formed, normally a sacrificial
5 oxide layer is grown on the walls of the trench and then removed. An oxide lining may then be formed on the walls of the trench. This stage yields a trench having a uniform oxide layer on its walls, with or without a hard mask on the top surface of the silicon.

One may then proceed to what is called the directed dielectric deposition,
10 which involves depositing more oxide on the bottom of the trench than on the sidewalls of the trench. There are then three choices. As shown in Fig. 16, a selective etchback can be performed, allowing thick oxide to remain at the bottom of the trench and removing the oxide from the sidewalls of the trench. As shown in Figs. 13-15 one can perform a "dipback" to remove the oxide layer from the
15 sidewalls of the trench. Finally, one can perform a selective oxidation, as shown in Figs. 17A and 18, wherein a polysilicon layer is formed at the bottom of the trench and then oxidized to form additional oxide at the bottom of the trench. The selective oxidation of a polysilicon layer can be performed instead of or in addition to the directed dielectric deposition.

20 At this stage of the process a trench has been fabricated with a thick oxide layer on the bottom. There may or may not be a "hard" mask on the top surface of the semiconductor. Next, a thin oxide layer is grown on the walls of the trench and the trench is filled with polysilicon. The polysilicon may be deposited as a single layer or it can be deposited as two layers with an etchback between the depositions.
25 Depositing the polysilicon in a two-stage process may be beneficial to the introduction of dopants into the "mesa" between the trenches, and to make a more lightly doped polysilicon layer available on the surface of the wafer to produce diodes, resistors, and other polysilicon devices.

Finally a glass layer is deposited and contact openings are formed in the
30 glass layer.

Figs. 13A-13N illustrate a process using the oxide "dipback" method. The process starts with an epi layer 262 formed on a substrate 260. A mask layer 264 is formed on the top surface of epi layer 262, with an opening where the trench is to be formed. Mask layer 264 may be photoresist or some other material and may be formed on top of an oxide layer 262. A trench 268 is then formed using conventional processes, as shown in Fig. 13A.

In Fig. 13B a sacrificial oxide layer 270 has been formed on the surface of the trench. Sacrificial oxide layer 270 is then removed, as shown in Fig. 13C. Sacrificial oxide layer 270 could be from 100 Å to 1000 Å thick; typically, it would be in the range of 300 Å thick. It can be formed by heating the structure at 800° C to 1100° C for 10 minutes to five hours in an oxidizing ambient. The ambient could be either oxygen or it could be oxygen and hydrogen. If the ambient is a combination of oxygen and hydrogen, it is considered a "wet" oxidation because the reaction would produce water vapor and this would affect the consistency and growth rate of the oxide.

Optionally, an oxide lining 272 is then formed on the walls of trench 268. Lining 272 could have a thickness in the range of 100 Å TO 600Å. Lining 272 prevents the deposited oxides from contacting the silicon directly, which would have the potential for charged states, especially at the interface between the silicon and the deposited oxide.. Adding a clean oxide layer on the walls of the trench provides a reduced charge state.

As shown in Fig. 13E, an electric field is applied above the surface of epi layer 262, and dielectric ions are formed and directed downward into trench 268 by means of the electric field. Preferably, a plasma-enhanced chemical vapor deposition chamber is used for this process. The electric field accelerates the dielectric ions downward so that they preferentially deposit on horizontal surfaces, including the bottom of trench 268. The chemical vapor deposition of oxide involves a gaseous chemical reaction of oxygen and silane, dichlorosilane, or silicon tetrachloride. The source of oxygen is typically nitreous oxide, and silane is

typically the silicon source. Plasma-enhanced chemical vapor deposition machines are available from such companies as Novellus Systems and Applied Materials.

Another method to achieve a directional deposition is to sputter a oxide film from an oxide-coated target onto the wafer. Since sputtering is a momentum transfer process, the deposition occurs in a straight line.

The result of this process is shown in Fig. 13F, where an oxide layer 270 has been formed inside and outside the trench 268. Note that oxide layer 270 is thicker at the bottom of trench 268 than on the sidewalls of trench 268. It is also thicker on the flat surfaces of epi layer 262. Processes other than chemical vapor deposition, such as sputtering, could also be used to produce oxide layer 270.

Layer 270 could be formed of materials other than oxide, such as phosphorus-doped glass or boron phosphorus silicon glass. It could also consist of other materials having a low dielectric constant K, such as polymers or polyimide. Air bubbles could be incorporated in layer 270 to reduce its dielectric constant.

In Fig. 13G, oxide layer 270 has been etched back or dipped back to remove the portions on the sidewalls of trench 268. A bottom portion 274 of oxide layer 270 remains at the bottom of trench 268. As shown in Fig. 13H, the structure is then heated to form a thin oxide layer 276 on the sidewalls of trench 268. A polysilicon layer 278 is then deposited to fill trench 268 and overflow the top surface of the structure. This is shown in Fig. 13I.

As shown in Fig. 13J, polysilicon layer 278 is then etched back until it is roughly coplanar with the top surface of epi layer 262. Next, the portions of oxide layer 270 on the surface of epi layer 262 are removed, taking care not to etch too much of the oxide layer 276 on the sidewalls of the trench. The result of this step is shown in Fig. 13K. Avoiding the removal of oxide layer 276 is best performed by having polysilicon layer 278 protrude slightly above the oxide layer 276. In Fig. 13L, the entire top surface of the structure, including the top surface of polysilicon layer 278, has been oxidized to form an oxide layer 280.